Statistical Mechanical Analysis of Compressed Sensing Utilizing Correlated Compression Matrix

Koujin Takeda

Department of Computational Intelligence and
Systems Science
Tokyo Institute of Technology
Yokohama 226-8502, Japan
E-mail: takeda@sp.dis.titech.ac.jp

Yoshiyuki Kabashima

Department of Computational Intelligence and
Systems Science
Tokyo Institute of Technology
Yokohama 226-8502, Japan
E-mail: kaba@dis.titech.ac.jp

Abstract—We investigate a reconstruction limit of compressed sensing for a reconstruction scheme based on the L_1-norm minimization utilizing a correlated compression matrix with a statistical mechanics method. We focus on the compression matrix modeled as the Kronecker-type random matrix studied in research on multiple-input multiple-output wireless communication systems. We found that strong one-dimensional correlations between expansion bases of original information slightly degrade reconstruction performance.

I. INTRODUCTION

A novel approach of data compression, termed compressed sensing (CS), has recently been drawing great attention. The central assumption of CS is the sparsity of original information, which seems plausible for many real world signals. For exploiting this property, much effort has been paid in both research directions of theory and application [1]–[3].

The basic idea of CS is summarized in the following linear equation:

$$y = Fx^0. \tag{1}$$

(Throughout this article vectors and matrices are denoted in bold letters). $x^0 \in \mathbb{R}^N$ denotes N-dimensional coefficient vector of the original information $f \in \mathbb{R}^N$ expanded over the basis $\phi_i \in \mathbb{R}^N$ ($1 \le i \le N$), namely $f = \sum_i x_i^0 \phi_i$, and $y \in \mathbb{R}^P$ is a P-dimensional vector, which describes compressed information available from P times observations with observation vector ψ_i, $y_i = f \cdot \psi_i$ ($1 \le i \le P$). F is a P-by-N compression matrix whose element is given by $F_{ij} = \psi_i \cdot \phi_j$. In this article the compression matrix F is regarded as random, which means that we are dealing with random expansion bases and random observation vectors. The compression rate is defined by $\alpha \equiv P/N < 1$. The original coefficient x^0 is sparse and modeled by the distribution,

$$P(x_i^0) = (1 - \rho)\delta(x_i^0) + \rho \exp\left(-(x_i^0)^2/2\right)/\sqrt{2\pi}, \tag{2}$$

that is, ρ represents the density of non-zero coefficients. Under the above setting, L_1-norm minimization offers an appropriate feasible algorithm for reconstruction of the original coefficient (termed L_1-norm reconstruction),

$$\text{minimize } \|x\|_1 \quad \text{subject to } y(= Fx^0) = Fx, \tag{3}$$

where $\|x\|_p = \lim_{\epsilon \to +0} \sum_i |x_i|^{p+\epsilon}$. The remaining problem is whether the solution x coincides with the original coefficient x_0. We can expect that below a certain critical value of the compression rate α_c, the original coefficient x^0 cannot be reproduced even if we make use of the L_1-norm reconstruction. The aim of this article is to evaluate this critical value α_c in the limit of $P, N \to \infty$ (and $\alpha = \text{const.}$) utilizing a statistical mechanics method.

By the way, this problem is quite similar to the performance evaluation problem of linear vector channels in wireless communication, and accordingly the analysis scheme with the statistical mechanical approach for code-division multiple-access (CDMA) or multiple-input multiple-output (MIMO) communication [4]–[6] can be applied in the limit $P, N \to \infty$. Kabashima et al. have already investigated the performance of the L_1-norm reconstruction (to be precise general L_p-norm, though $p \le 1$ is the reconstructable case) using a statistical mechanical method in a basic scenario in [7], where F is composed of independently and identically distributed (i.i.d.) random variables, and evaluated the reconstruction limit α_c. The evaluation value accords with the one that has been assessed in [8], [9] using combinatorial geometry methods. (As a related study, noisy compressed sensing was investigated using replica method in [10], for which perfect reconstruction is not possible as long as the noise intensity is not negligible).

In this article, as a second step of the investigation, we consider a more advanced case in which F is provided as

$$F = \sqrt{R_r}\,\Xi\,\sqrt{R_t}, \tag{4}$$

and our goal is to evaluate critical value α_c for such F. Here, R_r and R_t are a P- and an N-dimensional square symmetric matrix, respectively. The square root of a square matrix A is defined as $A = \sqrt{A}^T\sqrt{A}$. Ξ is a random P-by-N rectangular matrix whose elements are i.i.d. Gaussian random variables of zero mean and variance N^{-1}. This random matrix Ξ effectively implies a situation in that the expansion bases and the observation vectors are statistically uncorrelated. In this modeling, the matrices R_t and R_r represent the correlations among the expansion bases ϕ and those among the observation vectors ψ, respectively. Random matrix of this type, F, is known as the channel matrix in the Kronecker model of the

MIMO communication system, whose performance is investigated by Hatabu et al. [6] with a statistical mechanical scheme. Accordingly, by application of this method it is expected that the reconstruction limit of the L_p-norm reconstruction can also be estimated. In the subsequent sections we explain the details of the analysis.

II. REPLICA ANALYSIS

In this section we describe the outline of the analysis. As we mentioned, the analysis is based on that for the Kronecker channel in the MIMO communication system [6], and the details of the analysis are also discussed in this work.

Following the discussions in [7], let us first define the cost function of the L_p-norm reconstruction using the quenched average of free energy, which is a standard technique for dealing with a random system in statistical mechanics,

$$C_p \equiv -\lim_{\beta\to\infty}\lim_{n\to 0}\frac{\partial}{\partial n}\lim_{N\to\infty}\frac{1}{\beta N}\ln[Z^n(\beta,\boldsymbol{y})]_{\boldsymbol{F},\boldsymbol{x}^0}, \quad (5)$$

where $[\cdot]_{\boldsymbol{F},\boldsymbol{x}^0}$ denotes the average over the random matrix $\boldsymbol{F}$ and the original coefficient $\boldsymbol{x}^0$ with the distribution (2). We also define the replicated partition function $Z^n(\beta,\boldsymbol{y})$ for $n \in \mathbb{N}$ as

$$Z^n(\beta,\boldsymbol{y}) \equiv \prod_{a=1}^{n}\int d\boldsymbol{x}^a \exp(-\beta\,\|\boldsymbol{x}^a\|_p)\delta(\boldsymbol{F}(\boldsymbol{x}^a-\boldsymbol{x}^0))$$

$$= \prod_{a=1}^{n}\int d\boldsymbol{x}^a \lim_{\tau\to+0}\frac{1}{(\sqrt{2\pi\tau})^{nP}}$$

$$\times \exp\left[-\sum_{a=1}^{p}\beta\,\|\boldsymbol{x}^a\|_p -\frac{1}{2\tau}\sum_{a=1}^{n}(\boldsymbol{x}^a-\boldsymbol{x}^0)^T\boldsymbol{F}^T\boldsymbol{F}(\boldsymbol{x}^a-\boldsymbol{x}^0)\right]. \quad (6)$$

From these expressions, we easily see that the cost function C_p is nothing but the minimized norm with the constraint $\boldsymbol{y} \equiv \boldsymbol{F}\boldsymbol{x} = \boldsymbol{F}\boldsymbol{x}^0$ for given $\boldsymbol{y}$. After performing the average over the random matrix $\boldsymbol{\Xi}$ in $\boldsymbol{F}$ we have

$$\int d\boldsymbol{F}\int d\boldsymbol{x}^0 \prod_{a=1}^{n}\int d\boldsymbol{x}^a \lim_{\tau\to+0}\frac{1}{(\sqrt{2\pi\tau})^{nP}}$$

$$\times \exp\left[-\frac{1}{2\tau}\sum_{a=1}^{n}(\boldsymbol{x}^a-\boldsymbol{x}^0)^T\boldsymbol{F}^T\boldsymbol{F}(\boldsymbol{x}^a-\boldsymbol{x}^0)-\sum_{a=1}^{n}\beta\,\|\boldsymbol{x}^a\|_p\right]$$

$$= \int d\boldsymbol{x}^0 \lim_{\tau\to+0}\frac{1}{(\sqrt{2\pi\tau})^{nP}}$$

$$\times \int d\boldsymbol{Q}\exp\left[N\operatorname{Tr}G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}\left(-\frac{1}{\tau}\boldsymbol{S}\right)+\ln\Pi^{(n)}(\boldsymbol{Q})\right], \quad (7)$$

where $(\boldsymbol{S})_{ab} \equiv Q_{ab}-Q_{a0}-Q_{0b}+Q_{00}$ and $\boldsymbol{Q}$ is an n-dimensional matrix defined by the constraint,

$$\Pi^{(n)}(\boldsymbol{Q}) \equiv \prod_{a=1}^{n}\int d\boldsymbol{x}^a\left\{\prod_{a=1}^{n}\delta(\boldsymbol{x}^{aT}\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}^a-NQ_{aa})\right\}$$

$$\times\left\{\prod_{a<b}^{n}\delta(\boldsymbol{x}^{aT}\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}^b-NQ_{ab})\right\}\left\{\prod_{a=1}^{n}\delta(\boldsymbol{x}^{aT}\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}^0-NQ_{a0})\right\}$$

$$\times\{\delta(\boldsymbol{x}^{0T}\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}^0-NQ_{00})\}\exp\left(-\sum_{a=1}^{n}\beta\,\|\boldsymbol{x}^a\|_p\right). \quad (8)$$

The function $G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}$ is defined as

$$G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}(\boldsymbol{A}) \equiv -\frac{\alpha}{2}\int d\lambda\rho_{\boldsymbol{R}_{\mathrm{r}}}(\lambda)\ln\left(\boldsymbol{I}-\frac{\lambda}{\alpha}\boldsymbol{A}\right). \quad (9)$$

The function $\rho_{\boldsymbol{R}_{\mathrm{r}}}(\lambda)$ in the definition of $G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}$ is the eigenvalue distribution of the matrix $\boldsymbol{R}_{\mathrm{r}}$.

Assuming replica symmetry, let $q = Q_{ab}$ ($a \neq b$), $Q = Q_{aa}$, $m = Q_{a0}$, and $u = Q_{00}$. Here, u is defined as $u \equiv N^{-1}\int\prod_i dx_i^0 P(x^0)\boldsymbol{x}^{0T}\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}^0 = \frac{\rho}{N}\operatorname{Tr}\boldsymbol{R}_{\mathrm{t}}$. From these assumptions it follows that $S_{aa} = Q - 2m + u$ and $S_{ab} = q - 2m + u$ ($a \neq b$). By diagonalization of the matrix $\boldsymbol{S}$, we can evaluate the $G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}$ -dependent part,

$$\exp\left[N\operatorname{Tr}G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}\left(-\frac{1}{\tau}\boldsymbol{S}\right)\right]$$

$$= \exp\left[N\left\{G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}\left(-\frac{Q-q}{\tau}\right)\right.\right.$$

$$-\frac{n(q-2m+u)}{\tau}G'_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}\left(-\frac{Q-q}{\tau}\right)+O(n^2)$$

$$\left.\left.+(n-1)G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}\left(-\frac{Q-q}{\tau}\right)\right\}\right]. \quad (10)$$

From the saddle-point method, the $\boldsymbol{x}$-dependent part, including the L_p-norm and the constraint, is expressed as,

$$\Pi^{(n)}(\boldsymbol{Q})$$

$$\equiv \operatorname*{Extr}_{\tilde{Q},\tilde{q},\tilde{m}}\left(\exp\left\{-NnQ\tilde{Q}-N\frac{n(n-1)}{2}q\tilde{q}-Nnm\tilde{m}\right\}\right.$$

$$\times\int D\tilde{\boldsymbol{z}}\left(\int d\boldsymbol{x}\exp\left[\left(\tilde{Q}-\frac{\tilde{q}}{2}\right)\boldsymbol{x}^T\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}\right.\right.$$

$$\left.\left.\left.+\boldsymbol{x}^T\sqrt{\boldsymbol{R}_{\mathrm{t}}}^T\left(\tilde{m}\sqrt{\boldsymbol{R}_{\mathrm{t}}}\boldsymbol{x}^0+\sqrt{\tilde{q}}\tilde{\boldsymbol{z}}\right)-\beta\,\|\boldsymbol{x}\|_p\right]\right)^n\right), \quad (11)$$

where the interaction between replicas is removed by introducing auxiliary variable $\boldsymbol{z}$ (Hubbard-Stratonovich transformation). For simpler expression of C_p, let us define new variables $\widehat{m} \equiv \beta^{-1}\tilde{m}, \widehat{\chi} \equiv \beta^{-2}\tilde{q}, \chi \equiv \beta(Q-q), \widehat{Q} \equiv \beta^{-1}(-2\tilde{Q}+\tilde{q})$ and the function

$$\phi_p(\boldsymbol{h},\widehat{Q})\equiv\frac{1}{N}\lim_{\epsilon\to+0}\min_{\boldsymbol{x}}\left\{\frac{\widehat{Q}}{2}\boldsymbol{x}^T\boldsymbol{R}_{\mathrm{t}}\boldsymbol{x}-\boldsymbol{h}^T\sqrt{\boldsymbol{R}_{\mathrm{t}}}\boldsymbol{x}+\|\boldsymbol{x}\|_{p+\epsilon}\right\}. \quad (12)$$

For $\beta \to \infty$, the $\boldsymbol{x}^0$-dependent part is rewritten as

$$\prod_i\int P(x_i^0)dx_i^0\int D\tilde{\boldsymbol{z}}\exp\left(-\beta Nn\phi_p(\widehat{m}\sqrt{\boldsymbol{R}_{\mathrm{t}}}\boldsymbol{x}^0+\sqrt{\widehat{\chi}}\tilde{\boldsymbol{z}},\widehat{Q})\right). \quad (13)$$

Combining these results and inserting the expression of the function $G_{\boldsymbol{\Xi}^T\boldsymbol{R}_{\mathrm{r}}\boldsymbol{\Xi}}$ and its derivative in the limit $\tau \to +0$, we have the final expression of the cost function

$$C_p = \operatorname*{Extr}_{q,m,\chi,\widehat{Q},\widehat{m},\widehat{\chi}}\left(\frac{\alpha(q-2m+u)}{2\chi}+\left(\frac{\chi\widehat{\chi}}{2}-\frac{q\widehat{Q}}{2}+m\widehat{m}\right)\right.$$

$$\left.+\left\{\prod_i\int dx_i^0 P(x_i^0)\int D\tilde{\boldsymbol{z}}\phi_p(\widehat{m}\sqrt{\boldsymbol{R}_{\mathrm{t}}}\boldsymbol{x}^0+\sqrt{\widehat{\chi}}\tilde{\boldsymbol{z}},\widehat{Q})\right\}\right). \quad (14)$$

The cost function depends only on the correlation matrix $\boldsymbol{R}_{\mathrm{t}}$ between expansion bases of the original information, and does not depend on the matrix $\boldsymbol{R}_{\mathrm{r}}$ between observation vectors, which indicates that the observation procedure is not essential for CS. This seems reasonable because sparsity of the original coefficient is significant and observation is not for CS. Accordingly, we must concentrate only on the effect of the correlation matrix $\boldsymbol{R}_{\mathrm{t}}$ on the L_p-norm reconstruction.

As mentioned above, C_p is nothing but the minimized L_p-norm, and the expression Eq. (14) tells us that the minimized L_p-norm is given by the solution of the extremization problem. The remaining problem is whether the original coefficient is correctly reconstructed typically from the solution of the extremization problem. Remembering the fact that $\boldsymbol{x}$ is the result of the reconstruction and $\boldsymbol{x}^0$ is the original coefficient, from Eq.(8) $q = m(= u)$ must hold when the reconstruction is successful. Therefore, the scheme for finding the reconstruction limit α_c is as follows: vary the parameter α (compression rate) and ρ (density of non-zero coefficients), then solve the extremization problem, and examine whether the solution satisfies $q = m(= u)$.

We have completed the replica analysis as above, and the cost function (14) we obtained describes the information of the L_p-norm reconstruction for arbitrary correlation matrices $\boldsymbol{R}_{\mathrm{r}}$ and $\boldsymbol{R}_{\mathrm{t}}$. (As you see $\boldsymbol{R}_{\mathrm{r}}$ will eventually become irrelevant). However, one problem remains: the cost function C_p includes the function $\phi_p(\boldsymbol{h}, \widehat{Q})$, which is defined by the minimization problem with N variables, whose expression is not so simple (For L_1-norm this problem can be solved numerically in principle because the minimization function is unimodal). Fortunately, for certain classes of $\boldsymbol{R}_{\mathrm{t}}$, this minimization problem can be expressed relatively simply, which allows us to evaluate the reconstruction limit in a tractable manner. In the case without correlation $\boldsymbol{R}_{\mathrm{t}} = \boldsymbol{I}$, we see that the minimization problem with N variables is reduced to the problem with a single variable, and the cost function changes to the one obtained in [7]. In the following section, we will give another simple but nontrivial example, for which the minimization problem is numerically tractable.

III. EXAMPLE: ADJACENT CORRELATION

Let us consider that the correlation matrix $\boldsymbol{R}_{\mathrm{t}}$ has a tridiagonal form, as discussed in [6], in the context of the MIMO communication system, defined by

$$
\boldsymbol{R}_{\mathrm{t}} = \begin{pmatrix} 1 & r & 0 & \dots & r \\ r & 1 & r & \dots & 0 \\ 0 & r & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r & 0 & 0 & \dots & 1 \end{pmatrix}. \quad (15)
$$

This corresponds to the case that only adjacent matrix components (or adjacent expansion bases) have correlation. This matrix can be decomposed as $\boldsymbol{R}_{\mathrm{t}} = \sqrt{\boldsymbol{R}_{\mathrm{t}}}^T \sqrt{\boldsymbol{R}_{\mathrm{t}}}$ (Cholesky decomposition with boundary term), where

$$
\sqrt{\boldsymbol{R}_{\mathrm{t}}} = \begin{pmatrix} l_+ & l_- & 0 & \dots & 0 \\ 0 & l_+ & l_- & \dots & 0 \\ 0 & 0 & l_+ & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_- & 0 & 0 & \dots & l_+ \end{pmatrix}. \quad (16)
$$

Here $l_\pm \equiv (\sqrt{1+2r} \pm \sqrt{1-2r})/2$. In what follows we focus on the L_1-norm minimization. In the present case $u = \mathrm{Tr}\,\boldsymbol{R}_{\mathrm{t}}/N = \rho$ holds, and the cost function is rewritten as

$$
C_1 = \underset{q,m,\chi,\widehat{Q},\widehat{m},\widehat{\chi}}{\mathrm{Extr}} \left(\left\{ \frac{\alpha(q-2m+\rho)}{2\chi} + \left(\frac{\chi\widehat{\chi}}{2} - \frac{q\widehat{Q}}{2} + m\widehat{m} \right) \right. \right.
$$
$$
\left. \left. + \prod_i \int P(x_i^0) dx_i^0 \int \prod_j D\tilde{z}_j \phi_1(\widehat{\boldsymbol{h}}, \widehat{Q}) \right\} \right), \quad (17)
$$

where $\widehat{h}_i \equiv \widehat{m}(l_+ x_i^0 + l_- x_{i+1}^0) + \sqrt{\widehat{\chi}}\tilde{z}_i$ and x^0 is defined periodically as $x_{N+1}^0 = x_1^0$. (For simplicity we denote $\widehat{\boldsymbol{h}} \equiv \{\widehat{h}_1, \cdots, \widehat{h}_N\}$). The function $\phi_1(\widehat{\boldsymbol{h}}, \widehat{Q})$ in the cost function C_1 can be transformed as

$$
\phi_1(\widehat{\boldsymbol{h}}, \widehat{Q}) = \frac{1}{N}\min_{\boldsymbol{x}} \left\{ \frac{\widehat{Q}}{2}\boldsymbol{x}^T \boldsymbol{R}_{\mathrm{t}} \boldsymbol{x} - \widehat{\boldsymbol{h}}^T \sqrt{\boldsymbol{R}_{\mathrm{t}}}\boldsymbol{x} + \|\boldsymbol{x}\|_1 \right\}
$$
$$
= \frac{1}{N} \left\{ \frac{\widehat{Q}}{2}\sum_i (x_i^*)^2 + \widehat{Q}r \sum_i x_i^* x_{i+1}^* - \widehat{m}\sum_i x_i^* x_i^0 \right.
$$
$$
\left. -\widehat{m}r \sum_i x_i^* x_{i+1}^0 - \sqrt{\widehat{\chi}}\sum_i \tilde{z}_i(l_+ x_i^* + l_- x_{i+1}^*) + \sum_i |x_i^*| \right\}. \quad (18)
$$

The variables x, h are also defined periodically, $x_0 = x_N$, $x_{N+1} = x_1$, $h_0 = h_N$, and $h_{N+1} = h_1$. x_i^* is given by the solution of the minimization problem, namely for each i

$$
\frac{\partial}{\partial x_i^*}\phi_1(\boldsymbol{h}, \widehat{Q}) = (\widehat{Q}x_i^* - \widehat{m}x_i^0) + r(\widehat{Q}x_{i-1}^* - \widehat{m}x_{i-1}^0)
$$
$$
+ r(\widehat{Q}x_{i+1}^* - \widehat{m}x_{i+1}^0) - \sqrt{\widehat{\chi}}(l_+ \tilde{z}_i + l_- \tilde{z}_{i-1}) + \mathrm{sgn}(x_i^*)
$$
$$
= 0 \quad (19)
$$

is satisfied. As seen above, the minimization problem for each i includes only variables with three sequential indices $i-1, i$, and $i+1$, which indicates that the minimization problem is on a one-dimensional chain. It should be noted that the minimization function is unimodal, and sequential minimization for each variable enables us to find the minimum when we try to search it numerically. The computational cost of this procedure is $O(N)$ and feasible.

The extremization condition of the cost function can be expressed using the solution of the minimization problem,

denoted by x^*,

$$\widehat{Q} = \widehat{m} = \frac{\alpha}{\chi}, \qquad \widehat{\chi} = \frac{\alpha(q - 2m + \rho)}{\chi^2},$$

$$q = \frac{1}{N} \prod_i \int D\tilde{z}_i dx_i^0 P(x_i^0) \sum_i x_i^* (x_i^* + rx_{i-1}^* + rx_{i+1}^*),$$

$$m = \frac{1}{N} \prod_i \int D\tilde{z}_i dx_i^0 P(x_i^0) \sum_j x_j^* (x_j^0 + rx_{j-1}^0 + rx_{j+1}^0),$$

$$\chi = \frac{1}{\sqrt{\widehat{\chi}} N} \prod_i \int D\tilde{z}_i dx_i^0 P(x_i^0), \sum_j \tilde{z}_j (l_+ x_j^* + l_- x_{j+1}^*).$$
(20)

Remember that x_i^* depends on $\widehat{Q}, \widehat{m}, \sqrt{\widehat{\chi}}, \tilde{z}_i, x_i^0$ through Eq. (19).

The next issue is the reconstruction limit as discussed in [7]. As mentioned before, if the L_1-norm reconstruction works successfully, $q = m(= \rho)$ holds and the right hand side of the equation for $\widehat{\chi}$ in Eq. (20) vanishes. The extremization conditions for $\widehat{\chi}, q, m$ can be combined by using a new variable $\widehat{x}_i \equiv x_i^* - x_i^0$,

$$\widehat{\chi} = \frac{\alpha}{\chi^2 N} \prod_i \int D\tilde{z}_i dx_i^0 P(x_i^0) dx_i^0 \sum_j \widehat{x}_j (\widehat{x}_j + r\widehat{x}_{j-1} + r\widehat{x}_{j+1}).$$
(21)

From this expression it follows that $\widehat{x}_i = 0$ (namely $x_i^* = x_i^0$) is obtained from the extremization conditions as a solution of successful reconstruction. On the other hand, by inserting $\widehat{m} = \widehat{Q} = \alpha/\chi$ into Eq. (19),

$$\frac{\partial}{\partial x_i^*} \phi_1(\boldsymbol{h}, \widehat{Q}) = \frac{\alpha}{\chi} (\widehat{x}_i + r\widehat{x}_{i-1} + r\widehat{x}_{i+1})$$
$$- \sqrt{\widehat{\chi}}(l_+ \tilde{z}_i + l_- \tilde{z}_{i-1}) + \operatorname{sgn}(\widehat{x}_i + x_i^0) = 0. \quad (22)$$

This equation indicates that in the limit $\chi \to 0$ $\widehat{x}_i$ should vanish faster than $O(\chi)$ in order for the solution $\widehat{x}_i = 0$ to exist. (In this case $\widehat{\chi}$ is $O(1)$). From the insight above, we rescale the variable as $\widehat{x} \to (\chi/\alpha)\widehat{x}$. The remaining equations in terms of $\chi, \widehat{\chi}, \widehat{x}_i$ are

$$\widehat{\chi} = \frac{1}{\alpha N} \prod_i \int D\tilde{z}_i dx_i^0 P(x_i^0) \sum_j \widehat{x}_j (\widehat{x}_j + r\widehat{x}_{j-1} + r\widehat{x}_{j+1}),$$

$$\chi = \chi \left(\frac{1}{\alpha \sqrt{\widehat{\chi}} N} \prod_i \int D\tilde{z}_i dx_i^0 P(x_i^0) \sum_j \tilde{z}_j (l_+ \widehat{x}_j + l_- \widehat{x}_{j+1}) \right),$$

$$\frac{\partial}{\partial x_i^*} \phi_1(\boldsymbol{h}, \widehat{Q}) = (\widehat{x}_i + r\widehat{x}_{i-1} + r\widehat{x}_{i+1})$$
$$- \sqrt{\widehat{\chi}}(l_+ \tilde{z}_i + l_- \tilde{z}_{i-1}) + \operatorname{sgn}\left(\frac{\chi}{\alpha} \widehat{x}_i + x_i^0 \right) = 0. \quad (23)$$

The second equation has two solutions: $\chi = 0$ and the factor in the bracket is unity. The first solution $\chi = 0$ corresponds to successful reconstruction, $x_i = x_i^*$ (note that $\widehat{x}$ is rescaled), and the second (which satisfies $\chi \neq 0$) amounts to unsuccessful reconstruction. The equation for the threshold is obtained by inserting $\chi = 0$ to the second solution. In conjunction with the remaining two equations, we finally have the equations for the reconstruction limit (Note that in the evaluation of the

reconstruction limit, we must perform multiple integrals in the equations for χ and $\widehat{\chi}$, and we can use the Monte Carlo method in numerical evaluation).

Equation (23) is the main result. By the procedure mentioned above, we can estimate the reconstructions α_c as the function of ρ (density of non-zero coefficients) and r (correlation parameter in the compression matrix). For $\boldsymbol{R}_t = \boldsymbol{I}$ (without correlation), we can recover the reconstruction limit obtained in [7] directly from the expression of Eq. (23).

IV. EVALUATION OF RECONSTRUCTION LIMIT

For the adjacent correlation discussed above, we estimate the reconstruction limit by using the result of the replica analysis in Eq. (23). In Fig. 1, the dependences of reconstruction limit α_c on the density ρ are shown for uncorrelated ($r = 0$) and correlated ($r = 0.5$) cases. The difference between two results are very small over all region of ρ, which indicates the effect of adjacent correlation is very small. In Fig. 2 the dependence on the correlation parameter r is depicted for $\rho = 0.5$. In the case without correlation, Kabashima et al. obtained $\alpha_c = 0.8312...$ for $\rho = 0.5$ [7]. In the region of small r, we cannot observe the deviation of α_c from the uncorrelated case $r = 0$. On the other hand, for the strongly-correlated case $r = 0.5$, we observe a slight increase in α_c, which implies that strong correlation worsens the performance of the L_1-norm reconstruction. For $r = 0.5$, α_c is estimated as $\alpha_c = 0.84057(14)$, indicating that the performance falls about 1% from $r = 0$ in terms of the reconstruction limit.

For verification of the results from replica analysis, we also conducted a numerical experiment of the L_1-norm reconstruction. We used the convex optimization package for MATLAB developed in [11], [12]. The results are shown in Fig. 3. The dependence of α_c on the dimension of the original coefficient N is shown. For comparison with replica analysis, we also performed scaling analysis with quadratic function regression, and estimated the value of α_c for $N \to \infty$ limit. The results give $\alpha_c = 0.84017(28)$ for $N \to \infty$, which clearly indicates the increase in the value of α_c (or degradation of the reconstruction performance) as expected. The reconstruction limit estimated by extrapolation is very close to the one from replica analysis, which enforces the validity of the result from replica analysis.

V. CONCLUSIONS AND DISCUSSION

We investigated the performance of the L_p-norm (especially L_1-norm) reconstruction with a correlated compression matrix by replica analysis. We obtained the expression of the cost function for the Kronecker-type compression matrix for general correlation matrices $\boldsymbol{R}_t$ and $\boldsymbol{R}_r$.

The noteworthy issues in the results are summarized as follows. First, the cost function does not depend on the correlation matrix $\boldsymbol{R}_r$ describing the correlation between observation vectors, which indicates that the observation procedure is not significant in CS. This can also be understood from the fact that we can eliminate the correlation matrix $\boldsymbol{R}_r$ by redefinition (or rotation) of the random matrix $\boldsymbol{\Xi}$.
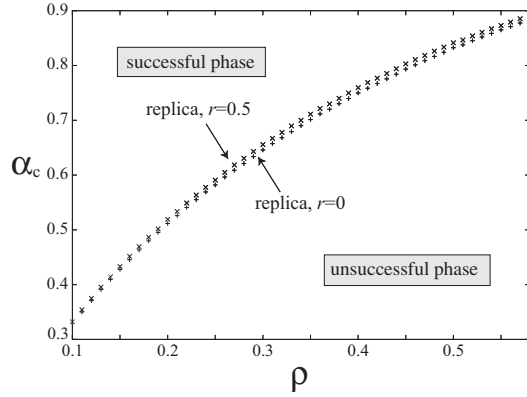
Fig. 1. Reconstruction limit α_c as a function of ρ. We compare the cases $r = 0.5$ ($\times$) and $r = 0$ (+, uncorrelated). We set the dimension of the original coefficient as $N = 10^6$. For the evaluation of the multiple integrals in Eq. (23), we perform the average over 100 samples of $\tilde{z}, x^0$. The difference between two results is very small.

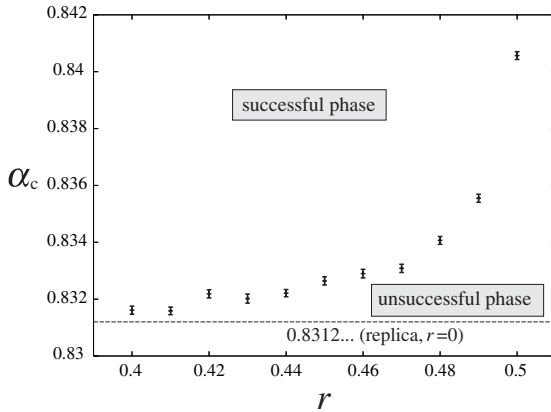Fig. 2. Dependence of reconstruction limit α_c on correlation parameter r. We set $\rho = 0.5$ and the dimension of the original coefficient as $N = 10^6$. For the evaluation of the multiple integrals in Eq. (23), we perform the average over 100 samples of $\tilde{z}, x^0$. In the vicinity of $r = 0.5$, the deviation of α_c from the value for $r = 0$ is clearly observed. (For $\rho = 0.5$, the reconstruction limit is estimated as $\alpha_c = 0.8312...$ [7].) For $r < 0.4$ the deviation is not visible and the result is not shown in this figure. At $r = 0.5$, the criticality is estimated as $\alpha_c = 0.84057(14)$.
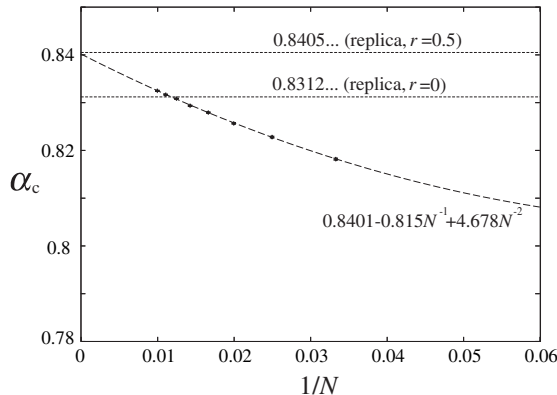

Fig. 3. Results of L_1-norm reconstruction experiment. We set $\rho = 0.5$ and $r = 0.5$. For each N (=dimension of the original information) we took the average over 10^5 samples. The broken line indicates the scaling by quadratic function regression. For $N \to \infty$, we clearly see the increase in the reconstruction limit from the case of no correlation ($\alpha_c = 0.8312...$ for $\rho = 0.5$ [7]), which is consistent with the result from replica analysis. The value at $N \to \infty$ from the extrapolation is $\alpha_c = 0.84017(28)$, which is very close to the value from replica analysis.

Second, we considered the correlated compression matrix, whereas the original coefficient is uncorrelated. However, our result can be reinterpreted as the case of a correlated original coefficient and uncorrelated expansion bases by redefinition (or rotation) of these quantities.

Finally, we found that the performance of the L_1-norm reconstruction is robust against the small correlation between adjacent expansion bases of the original coefficient. By incorporating the strong correlation, the reconstruction performance slightly falls. This is a quite natural result because the correlation between expansion bases implies the loss of the original information by redundant bases, which makes the reconstruction much more difficult. We should also keep in mind that we discussed the reconstruction limit only in the case that the correlation matrix is tridiagonal, and there is a possibility that another correlation matrix might highly degrade the performance (note that our analysis offers the result for general correlation), which will be for future work.

REFERENCES

[1] E. J. Candès, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inform. Theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.
[2] D. L. Donoho, "Compressed sensing," IEEE Trans. Inform. Theory, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.
[3] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," IEEE Trans. Inform. Theory, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.
[4] T. Tanaka, "A Statistical-Mechanical Approach to Large-System Analysis of CDMA Multiuser Detectors," IEEE Trans. Inform. Theory, vol. 48, no. 11, pp. 2888-2910, Nov. 2002.
[5] K. Takeda, S. Uda and Y. Kabashima, "Analysis of CDMA systems that are characterized by eigenvalue spectrum," Europhys. Lett., vol. 76, no. 6, pp. 1193-1199, Dec. 2006.
[6] H. Hatabu, K. Takeda and Y. Kabashima, "Statistical mechanical analysis of the Kronecker channel model for MIMO wireless communication," Phys. Rev. E, vol. 80, 061124, Dec. 2009.
[7] Y. Kabashima, T. Wadayama and T. Tanaka, "A typical reconstruction limit for compressed sensing based on L_p-norm minimization," J. Stat. Mech., L09003, Sep. 2009.
[8] D. L. Donoho, "High-Dimensional Centrally Symmetric Polytopes with Neighborliness Proportional to Dimension," Discrete Comput. Geom., vol. 35, no. 4, pp. 617-652, 2006.
[9] D. L. Donoho and J. Tanner, "Counting faces of randomly projector polytopes when the projection radically lowers dimension," J. Amer. Math. Soc., vol. 22, no. 1, pp. 1-53, Jan. 2009.
[10] S. Rangan, A. K. Fletcher and V. K. Goyal, "Asymptotic Analysis of MAP Estimation via the Replica Method and Compressed Sensing," NIPS 2009, pp. 1545-1553.
[11] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming (web page and software) 2009. http://stanford.edu/~boyd/cvx
[12] M. Grant and S. Boyd, Graph implementations for nonsmooth convex programs, Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, pp. 95-110, Lecture Notes in Control and Information Sciences, Springer 2008. http://stanford.edu/ boyd/graph_dcp.html